CLASSIFICATION OF SLEEP STAGES IN INFANTS: A NEURO FUZZY APPROACH

J. E. Heiss, C. M. Held, P. A. Estévez, C. A. Perez, C. A. Holzmann, J. P. Pérez

Department of Electrical Engineering, Universidad de Chile, Casilla 412-3, Santiago, Chile

Abstract - **An ANFIS based neuro-fuzzy system to classify sleep-waking states and stages in healthy infants has been developed. The classifier takes five input patterns identified from polysomnographic recordings on 20 s frames and assigns them to one out of five possible classes (WA, NREM-I, NREM-II, NREM-III&IV or REM). Eight polysomnographic recordings of healthy infants were studied, making a total of 3510 frames. Of these, four recordings were used for training, two for validation and two for testing. Results on the testing data achieved on average 88.2% of expert agreement in sleep-waking state-stage classification. These results were compared with the ones obtained using a multi-layer perceptron neural network (87.3%) and by applying the expert's rules for sleep classification (86.7%). The neuro-fuzzy approach also rendered fuzzy classification rules, which were analyzed and compared with the expert's rules.**

Keywords - **Sleep classification, sleep stages, ANFIS, neuro-fuzzy classifier, automated rule generation.**

I. Introduction

The sleep classification process is divided into three steps: data acquisition, pattern identification and sleep-waking state-stage classification. In the first step, several signals generated by bioelectrical and biomechanical activity of the infant's body are recorded by a polygraph, generating a large number of pages with graphical data. The pattern identification process is performed for each page. The expert determines the background predominant frequency range in the electroencephalogram (EEG) according to [1-3]; relevant for this study are the slow delta (SD) (0.5-2 Hz) and theta (TH) (3-7 Hz) frequency ranges. The EEG is also examined to detect sigma spindles (SS), which are in the 12-14 Hz range. The electrooculogram (EOG) and the electromyogram (EMG) are used to determine the presence of rapid eye movements (REMov) and muscular tone (MT), respectively. The polygraph records additional signals which the expert uses as context information, such as electrocardiogram (ECG), detection of body movements (BM), abdominal ventilatory movements, nostrils airflow, body temperatures and oxymetry.

The most basic division in sleep classification is to distinguish between wakefulness (WA) and being asleep. There are two sleep states called REM and Non-REM (NREM). NREM is subdivided in turn into four stages called NREM-I, NREM-II, NREM-III and NREM-IV. NREM-III and IV were considered as a single stage called NREM-III&IV in this study. The difference between the two is the threshold of SD presence.

To determine the sleep state or stage, the experts established certain rules, based on [2-3], that are shown in table I. However, sleep classification is not completely standardized and usually experts from different research centers have slightly different approaches. Even between expert co-workers there is usually less than 90% agreement in sleep classification [4].

TABLE I
EXPERT'S RULES FOR SLEEP-WAKING STATE-STAGE CLASSIFICATION
A: absent, P: present, X: irrelevant. A particular state or stage has to last at least one minute to be assessed as such.

Pattern	Sleep-Waking States & Stages				
	NREM-I	NREM-II	NREM-III&IV	REM	WA
REMov	A	A	A	P	P
TH	P	X	X	P	X
SD	A	A	P	A	A
SS	A	P	X	A	A
MT	X	X	X	A	P

The large amount of data, the complexity of the classification analysis and the variability among human experts are reasons to develop an automated sleep classification system. An evaluation of the computerized system ALICE 3 using fifty subjects [5] showed substantial differences between automated computer scoring and manually scored paper polysomnographies. A manual edition of the computer scoring enhanced agreement to 75.7% with the paper polysomnography scoring. In [1] a pattern identification system for sleep stage classification which emulates the way the expert searches for each of the five relevant patterns was implemented. A ganglionar lattice system performed the classification, achieving 84.9% of expert agreement, after manually removing several "noisy" pages from the database. Later on, in [6], the pattern detection algorithms were redesigned in order to enhance their robustness, and evaluated with an enlarged database using the expert's rules of table I. An 86.7% of correct classification was achieved for the testing set, which had no manual intervention and included "noisy" data.

In order to discover rules that may explain how the classification process should be performed and to find parameters that define the degree of presence or absence for a pattern, a neuro-fuzzy approach was chosen. The weight of each rule and the parameters of the membership functions were determined by supervised learning through an ANFIS [7] based neuro-fuzzy classifier (NFC) [8]. Non-relevant rules were eliminated by applying a pruning algorithm. The remaining rules were analyzed and compared with the expert's rules.

II. Methods

A. Data Acquisition

Eight continuous sleep recordings were obtained from infants between 6 and 13 months of age on a TECA lA97

Report Documentation Page

Report Date	Report Type	Dates Covered (from... to)
25OCT2001	N/A	-

Title and Subtitle	Contract Number
Classification of Sleep Stages in Infants: A Neuro Fuzzy Approach	
	Grant Number
	Program Element Number

Author(s)	Project Number
	Task Number
	Work Unit Number

Performing Organization Name(s) and Address(es)	Performing Organization Report Number
Department of Electrical Engineering, Universidad de Chile, Casilla 412-3, Santiago, Chile	

Sponsoring/Monitoring Agency Name(s) and Address(es)	Sponsor/Monitor's Acronym(s)
US Army Research, Development & Standardization Group (UK) PSC 802 Box 15 FPO AE 09499-1500	
	Sponsor/Monitor's Report Number(s)

Distribution/Availability Statement
Approved for public release, distribution unlimited

Supplementary Notes
Papers from the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, October 25-28, 2001, held in Istanbul, Turkey. See also ADM001351 for entire conference on cd-rom.

Abstract

Subject Terms

Report Classification	Classification of this page
unclassified	unclassified

Classification of Abstract	Limitation of Abstract
unclassified	UU

Number of Pages
4

18-channel polygraph connected as follows: five EEG channels with electrode placement adapted for infants from the international 10-20 system (FP1-C3, C3-O1, FP2-C4, C4-O2, and C3-C4); EOG for REMov detection; tonic chin and diaphragmatic EMGs; ECG; body movement detection of upper and lower limbs using piezo-electric crystal transducers; abdominal ventilatory movements, using a mercury strain gauge; and nostrils airflow, by means of a thermistor. All data were simultaneously recorded on paper and on digital means at a 250-Hz sampling rate. The digital data were collected on hard disk and then stored in laser media for off-line analysis. Infant behavior was also observed directly and noted on the polygraph paper. Depending on the polygraph settings, a page can last 20 or 30 seconds. The digital recordings were divided in 20 s frames, which represented one paper page in most cases.

B. Pattern Identification

The system described in [6] was applied to obtain a level of presence for each of the five relevant patterns. The pattern detection system outputs are either percentages of presence or quality indices of a given pattern per frame. The outputs are in the [0, 1] range.

The data set was divided into 4 records with 2067 frames for the training set, 2 records with 585 frames for the validation set and 2 records with 858 frames for the testing set. The training set was used to adjust the parameters with supervised learning in order to achieve over 80% agreement on the validation set, for each sleep-waking state-stage when using the expert's rules of table I. Two additional recordings were left for testing the system.

C. ANFIS based Neuro-Fuzzy Classifier

A neuro-fuzzy classifier (NFC) based on [7-8] was applied on the detected patterns to perform sleep-waking state-stage classification. Each of the five relevant patterns were associated to two fuzzy concepts, present and absent, with sigmoidal fuzzification functions.

$$1/1 + e^{-s(x-c)} , \qquad (1)$$

where s is the slope and c is the center of the sigmoid. The sign of s determines if the concept means present (+) or absent (-). Parameters s and c are determined through a training process, using the delta rule

$$\Delta W = -\mu \frac{\partial \varepsilon}{\partial W}, \qquad (2)$$

where ΔW is the adjustment for the parameter W, ε is the sum of the squared error, and μ is the learning rate. The weights of the linear combinations at layer 3 were also determined by supervised learning using the delta rule (2). The NFC architecture allowed us to implement a fuzzy classification system with differentiable fuzzification functions at layer 1 (in our case sigmoidal functions), including parameters that were trained using the delta rule
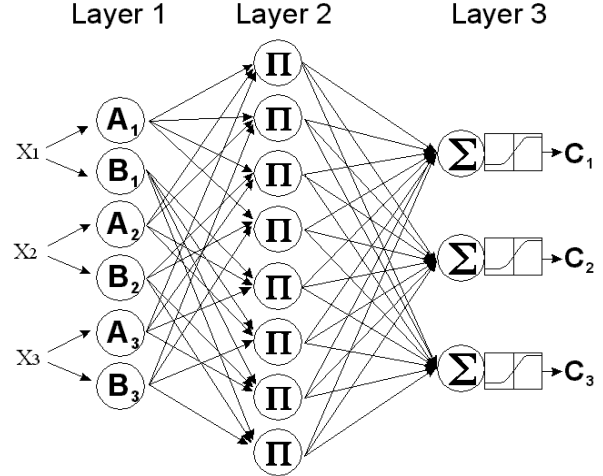

FIGURE 1. NEURO-FUZZY CLASSIFIER ARCHITECTURE.
Layer 1 is the fuzzification layer. Three input variables are shown here (X_1, X_2 and X_3), each with two associated fuzzy concepts (A_i and B_i). Layer 2 generates all the possible rules of the form IF X_1 is A_1 and X_2 is B_2 and X_3 is A_3, with a T-norm operator (Π), considering one fuzzy concept per input variable. The output of layer 2 is a strength parameter for each of the rules. Each node at layer 3 performs a linear combination of the rules and uses a sigmoidal function to determine the degree of belonging of the input pattern to each class (C_1, C_2, C_3).

(2) with the squared error as the objective function

$$\varepsilon = \sum_{j=1}^{m} \sum_{i=1}^{n} (d_i^j - o_i^j)^2 , \qquad (3)$$

where m is the number of output classes, n is the number of examples, d is the desired output (0 or 1) and o is the node output at layer 3.

The validation set was used to establish when to stop the learning algorithm, and consider the tuned parameters as final. The testing set was used to evaluate the performance of the tuned system with independent data. Nonlinear relations resulted after training between the NFC input and output spaces.

A simplified diagram of the NFC system architecture is shown in Fig. 1. A detailed explanation about the training process of an ANFIS network is given in [7]. The actual NFC model applied to the sleep problem had 5 inputs, each one with two associated fuzzy membership functions, and 5 output classes (WA, NREM-I, NREM-II, NREM-III&IV and REM). The combination of the fuzzy concepts of layer 1 produced a total of 32 rules at layer 2. The maximum output at layer 3 determined the class associated to each input vector. The weights at layer 3 were initialized with random values in the [0, 1] interval. The center c was initially determined at half of the maximum input from all the respective training examples and the parameter s was set at ± 5.

D. Postprocessing

In order to reduce the number of rules and thus produce a more expert-like set of fuzzy if-then rules, a pruning algorithm was implemented. For every output class, the average contribution of each rule was evaluated and a threshold of 0.01 was used to eliminate the least significant rules (the observed contribution values were always in the [0, 1] range). The last step of the classification process took into account that, according to expert criteria, every sleep-waking state-stage had to last at least one minute [2]. A state duration algorithm (SDA) was developed to ensure this condition [6].

In order to compare the performance of the system with a general classification method, a multi-layer perceptron (MLP) neural network with 5 input nodes, a hidden layer with 10 nodes and 5 output nodes was trained, using the same training, validation and testing sets as for the NFC.

III. RESULTS

Ten simulations with the ANFIS based NFC were performed and the test results were post-processed applying the SDA algorithm. The average results for the training, validation and test sets are summarized in table II. This table also shows the results of classifying these sets using a MLP neural network, and using the expert's rules of table I.

Only a few of the 32 rules survived after applying the pruning algorithm, for each of the 5 possible outputs (classes). As an example, the rules generated to classify NREM-I with the results of one of the ten simulations will be described in what follows. A similar analysis could be performed for all the other output classes. Fig. 2 shows the average contribution to the node output for the rules that were not pruned. Table III shows the surviving rules (R1, R2 and R3), with their respective fuzzy concepts associated to each rule (absent or present). Only the examples classified as NREM-I by the NFC were considered in the average calculation.
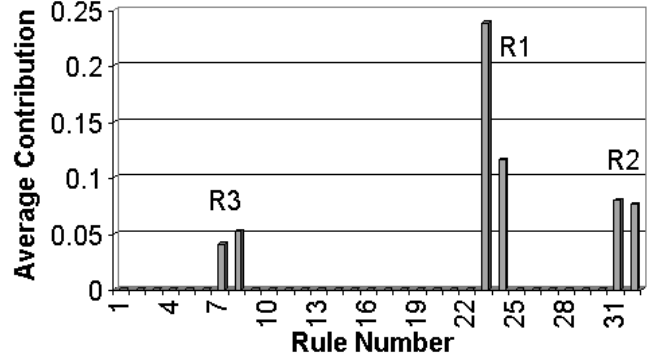

FIGURE 2. SYSTEM RULES FOR NREM-I

The average contribution of the surviving rules for NREM-I after pruning is shown. To simplify the analysis, the rules have been grouped in pairs (R1, R2 and R3), in accordance with their order of average contribution. The only difference in each pair of rules is the fuzzy concept associated to the MT input (presence or absence).

Table IV shows the system performance for classifying NREM-I after eliminating one of the three rules. Finally, Table V shows the relative activation frequency of rules R1, R2 and R3, as a function of the sleep-waking state previous to NREM-I. A rule was considered active when its contribution to the output was above 0.2 (in the [0,1] range).

TABLE III
FUZZY RULES GENERATED TO ASSESS STAGE NREM-I
The letters represent: A: absent, P: present, X: irrelevant. Absent and Present are fuzzy concepts defined by sigmoidal functions.

Pattern	Fuzzy Rules Generated		
	R1	R2	R3
REMov	A	A	A
TH	P	A	P
SD	A	A	P
SS	A	A	A
MT	X	X	X

TABLE IV
CLASSIFICATION PERFORMANCE DEGRADATION FOR STAGE NREM-I WHEN ONE OF THE RULES (R1, R2, R3) IS SUPRESSED

NREM-I Stage	Classification Performance			
Supressed Rule	None	R1	R2	R3
% of Correct Classification	82.4%	19.7%	65.0%	76.8%

TABLE V
RELATIVE ACTIVATION FREQUENCY FOR RULES R1,R2 AND R3 IN NREM-I

Previous State to NREM-I	R1	R2	R3
NREM	40.7%	70.0%	100%
REM OR WA	59.3%	30.0%	0%

TABLE II
PERFORMANCES OF THE IMPLEMENTED NFC, A MULTI LAYER PERCEPTRON NEURAL NETWORK (MLP) AND THE EXPERT'S RULES
The results show the over-all classification performance on a frame by frame basis except for the last column, which shows results on a one-minute basis after applying the State Duration Algorithm (SDA).

	Training	Validation	Test	Test with SDA
NFC	86.2 ± 0.1%	87.7 ± 0.2%	83.9 ± 0.4%	88.2± 0.5%
MLP	87.1 ± 0.7%	87.3 ± 0.4%	83.4 ± 0.6%	87.3 ± 0.9%
Expert's Rules	84.1%	87.2%	82.6%	86.7%

IV. DISCUSSION

The results of applying MLP and NFC (table II) were statistically non-different at a level of significance of 0.01 (t-test), for all data sets (training, validation and test). Both methods show an enhancement over applying the crisp expert's rules of table I. The last column of table II shows the results after applying the state duration algorithm, which improved the classification percentages because it eliminated isolated frames with different patterns. The same partition of sets used in [6] was maintained in order to perform meaningful comparisons between NFC, the crisp classifier and the MLP.

To evaluate the pruning algorithm, the results of the NFC applied to the training, validation and test set with and without pruning were compared, showing no statistically significant differences.

Fig. 2 and Tables III and IV show that there is a hierarchy among rules. R1 can be considered as the main rule while R2 and R3 are complementary rules; their combination made the system achieve a performance of over 80%. R1 matches exactly the expert's rule for NREM-I (table I). R2 and R3 are new discovered fuzzy rules.

Table V shows relative activation frequency for the surviving rules as a function of the preceding sleep state. It shows that R2 and R3 activate mainly within NREM sleep. These results suggests that NREM-I sleep within NREM may have different characteristics than NREM-I following WA or REM state. The rules R2 and R3 may help to identify this differences.

V. CONCLUSIONS

An ANFIS based neuro-fuzzy classifier with a pruning algorithm was implemented and applied to the classification of sleep–waking states-stages in infants, using the sleep pattern detection system of [6] to generate the inputs. Including artifacted pages, an average of 88.2% of expert agreement was achieved for testing data. As a result of the training process and pruning, rules and parameters that defined a fuzzy classification system were also determined. Analyzing the rules obtained for sleep stage NREM-I, it was found that the main rule matched the expert rule to classify NREM-I. Additional rules were discovered that complement the classification and may provide additional information about the characteristics of this sleep stage. This is a promissory result, and further research is needed in this topic.

Future work includes implementation of a clustering algorithm to determine the initial parameters of the system, training the system with a different objective function, such as the max-type error function described in [8], and evaluating the performance of different T-norms at layer 2

in Fig. 1. The development of a general methodology for rule discovery and interpretation is also of interest.

REFERENCES

[1] C.A. Holzmann, C.A. Perez, C.M. Held, M. San Martín, F. Pizarro, J.P. Pérez, M. Garrido, P. Peirano, "Expert System Classification of Sleep-Waking States in Infants", Medical & Biological Engineering & Computing, vol. 37, pp. 466-476, 1999.

[2] A. Rechtschaffen, A. Kales, "A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects", UCLA, Brain Research Institute/Brain Information Service, Los Angeles, USA, 1968.

[3] C. Guilleminault, M. Souquet, "Sleep States and Related Pathology", in R. Korobkin, and C. Guilleminault (Eds.), "Advances in Perinatal Neurology", Vol. 1, Spectrum, New York, USA, 1979.

[4] R. Norman, I. Pal, C. Stewart, J. Walsleben, D. Rappaport, "Interobserver Agreement Among Sleep Scorers From Different Centers in a Large Dataset", Sleep, Vol. 23, N° 7, pp. 901-908, 2000.

[5] D.P. White, T.J. Gibb, "Evaluation of a Computarized Polysomnographic System", Sleep, Vol. 21, N° 2, pp.188-196, 1998.

[6] P.A. Estévez, C.M. Held, C.A. Holzmann, C.A. Perez, J. Heiss, J.P. Pérez, M. Garrido, P. Peirano, "Polysomnographic Pattern Recognition for Automated Classification of Sleep-Waking States in Infants", Medical & Biological Engineering & Computing, 2001, unpublished.

[7] J.S.R. Jang, "ANFIS: Adaptive-Network-based Fuzzy Inference Systems," IEEE Trans. on Systems, Man, and Cybernetics, vol. 23, pp. 665-685, May 1993.

[8] J.S.R. Jang, C.T. Sun, "A Neuro-Fuzzy Classifier and its Applications", 2nd IEEE International Conference on Fuzzy Systems, Sn. Francisco, CA, USA, pp. 94-98, March 28–April 1, 1993.

Jaime Heiss
Área de Automática y Bioingeniería
Departamento de Ingeniería Eléctrica
Universidad de Chile
Av. Tupper 2007, Piso 3, Casilla 412-3, Santiago, Chile
E-mail: jheiss@cec.uchile.cl